Introduction

Computer Vision

Robby T. Tan

What is Computer Vision?

Human Visual System: What Do You *See*?

Definition

Computer vision is about how to make computers
work like human eyes

(**or beyond**).

Input and Output

Inputs:

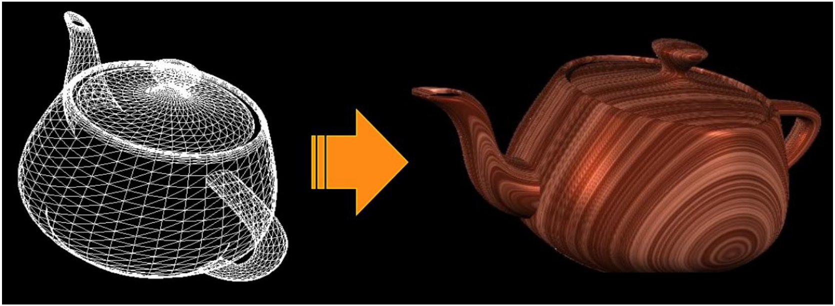
- Visual Data (primarily: 2D images or videos)

Outputs:

- Information of the visual contents of the input

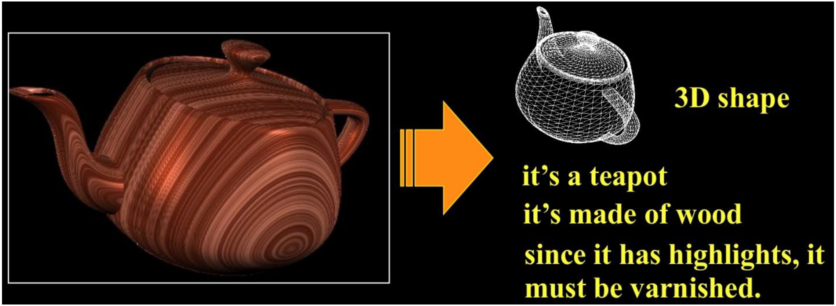
Related Fields: Differences

1. Computer Graphics
2. Image Processing
3. Artificial Intelligence



Computer Graphics: from mathematical models of a 3D world to 2D images

Computer Vision v.s. Computer Graphics



Computer Vision: from 2D images to the visual information

Image Processing

Image Processing: from an image to another image (which is visually better to a human observer)

Computer Vision and Image Processing

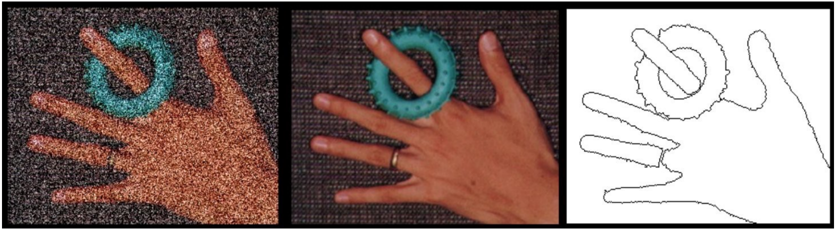


Image processing is part of low level computer vision.

- **Image processing**: given the left image, to produce the middle or the right image.
- **Computer vision**: given any image above, to identify the hand in the photograph.

Articifial Intelligence:

- A branch of computer science to extract information from all types of data (such as sound, text, visual, signal, etc.) through reasoning, deduction, planning, and learning.

Computer Vision:

- A branch of artificial intelligence focusing on visual data as the input.

Why Computer Vision?

Reasons of Studying Computer Vision

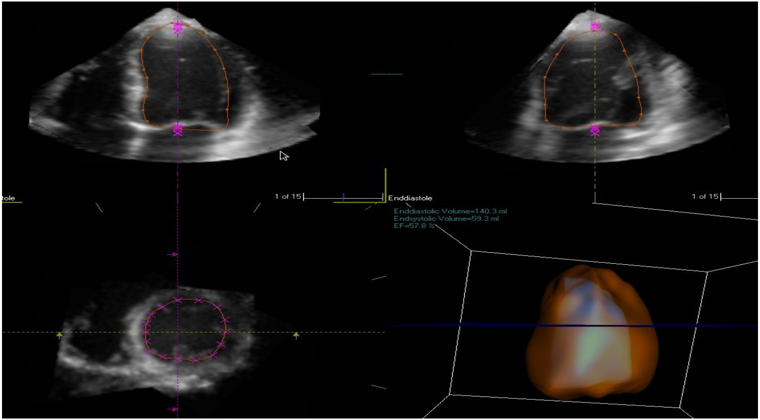
1. If you think your eyes are important, so is computer vision.
2. The potential applications are immense:
   - Game technology
   - Biomedical imaging
   - Robotics
   - Security/surveillance
   - Computer graphics
   - Human-computer interactions
   - Intelligent vehicle systems
   - Many more...
3. Challenging, but FUN, really really fun ;)

Self Driving Cars

Surveillance

Tracking and activity recognition help surveillance systems.

Medical Imaging



Real-time 3D echochardiography (cardiac/heart ultrasound):
www.ibme.ox.ac.uk/biomedia

Human Computer Interaction/Mixed Reality

Human interaction with mobile devices: the SixthSense (MIT)

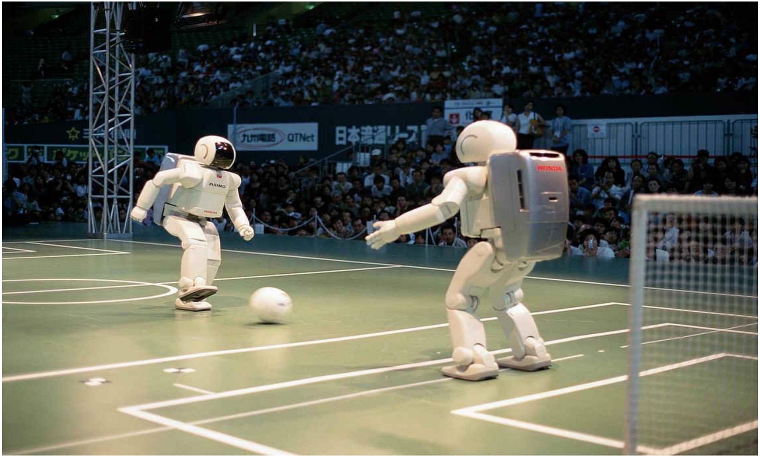
Computer Graphics

To realistically render a moving object from different angles requires the geometric (e.g. 3D data) and photometric properties (e.g. lighting).

Game Technology

To recognize the human poses in real time.

Robotics



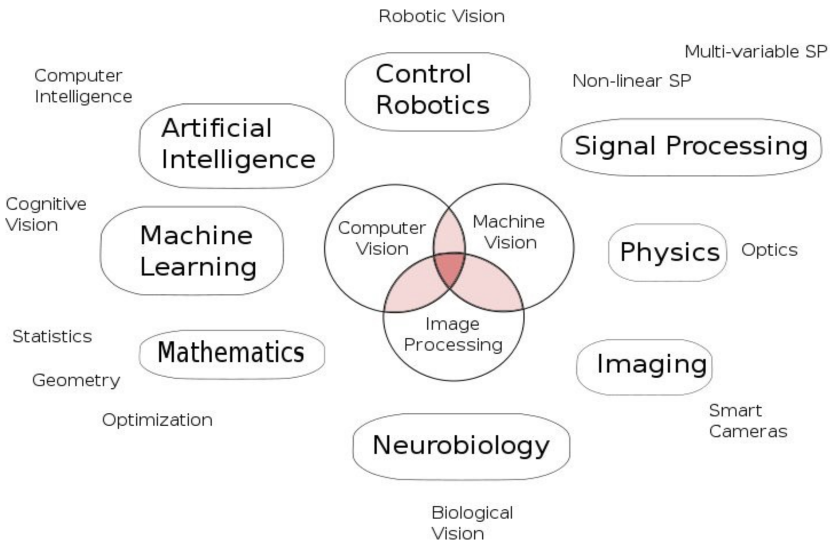
Robot vision is a direct application of computer vision.

Many Other Applications

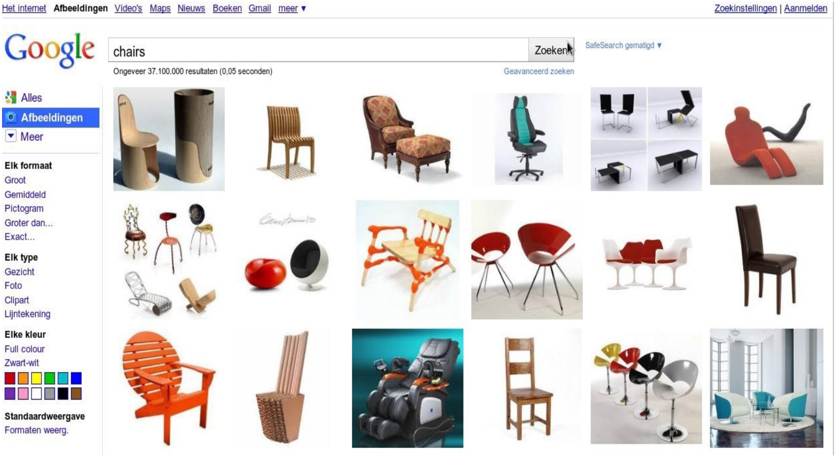
There are still many other applications of computer vision.

It is anticipated that computer vision systems
will soon become commonplace
and its technology will be applied for a broad range of products.

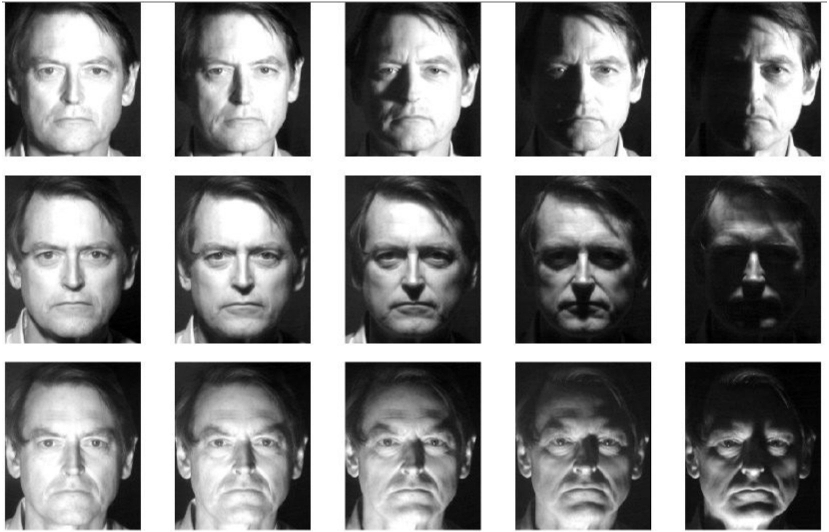
Related Fields



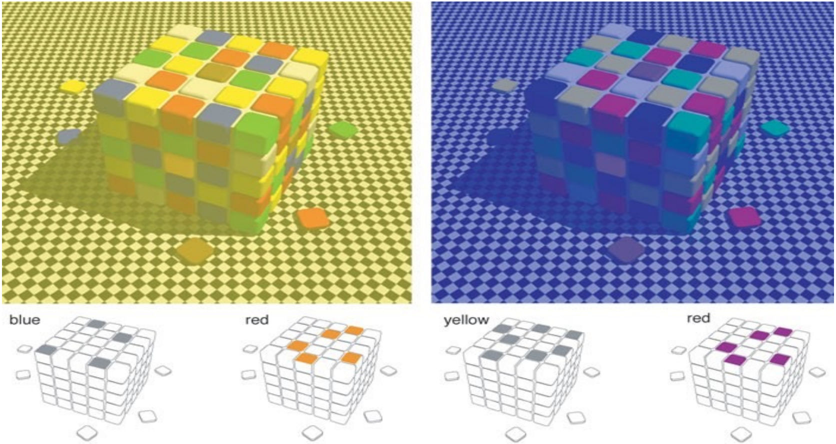
Challenge: Variation and Definition



Challenge: Lights

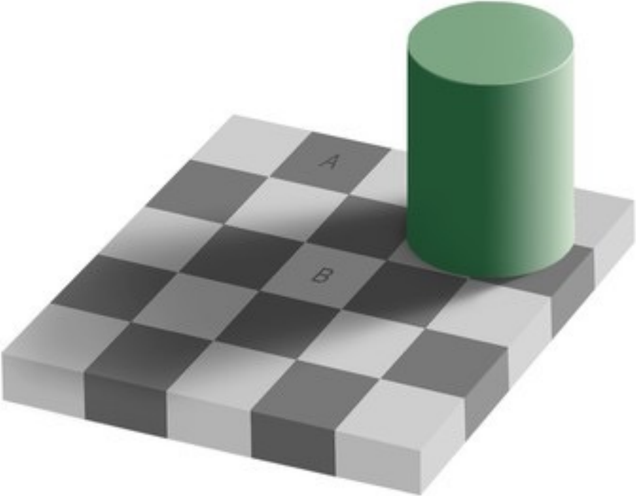


Challenge: Colors

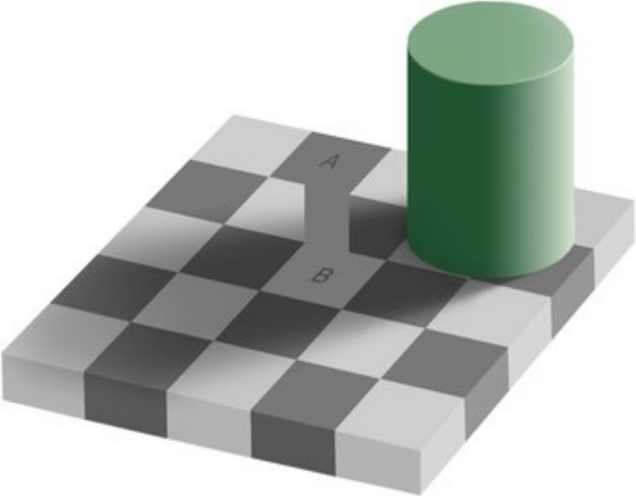


Don’t believe it? Check them yourself using an image editor!

Challenge: Shadow and Brightness



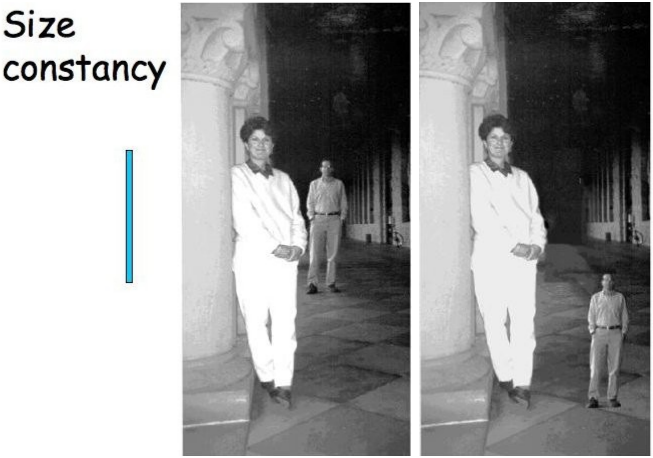
Challenge: Shadow and Brightness



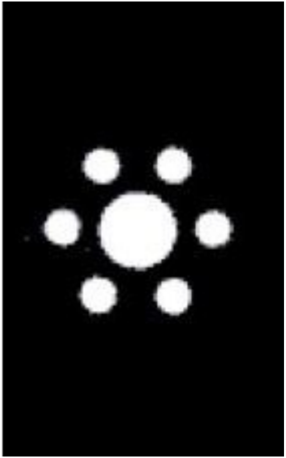
Challenge: Size

Size constancy

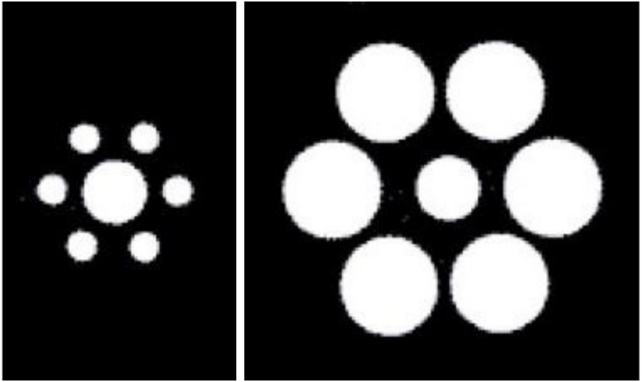
Challenge: Size



Challenge: Relative Size

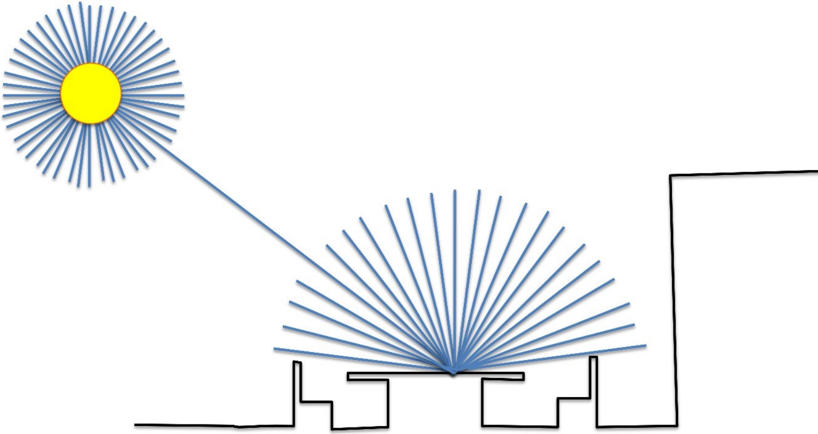


Challenge: Relative Size

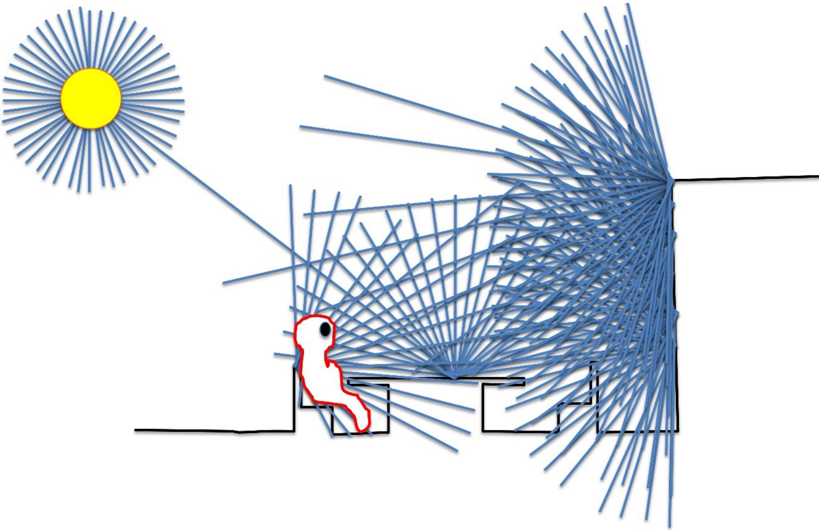


Challenge: Patterns

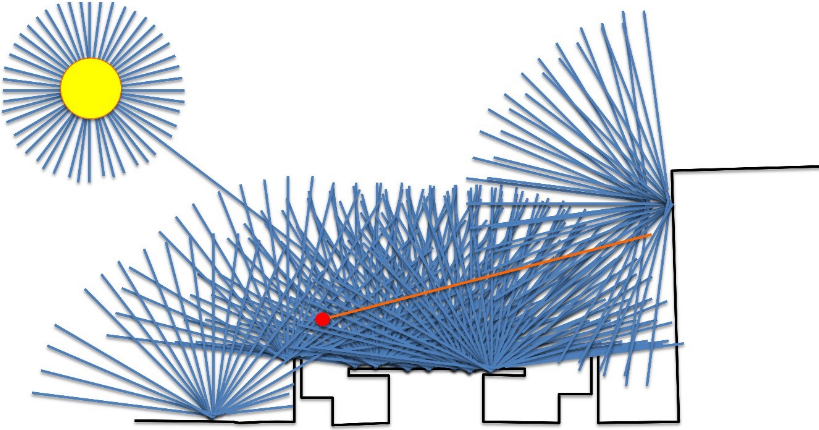
Challenge: Light Fields



Challenge: Light Fields



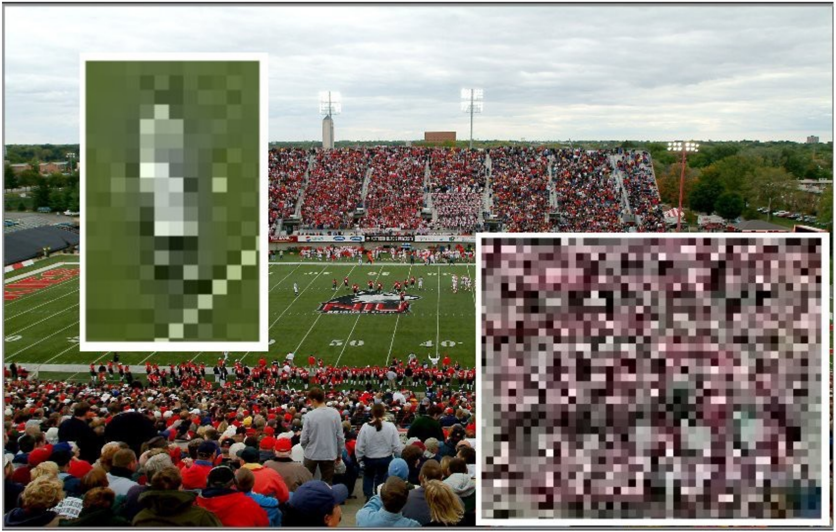
Challenge: Light Fields



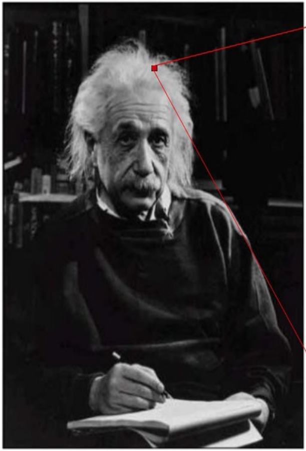
Challenging, but FUN, really FUN!

Challenging, but FUN, really FUN!

Challenging, but FUN, really FUN!



Our Input Data are Numbers!



123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
137	116	198	122	152	127	203	117	155	161	167	149	143	159	116	93
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
149	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
165	133	176	98	128	111	173	201	152	172	165	97	113	125	122	129
172	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
134	175	132	255	111	161	114	112	99	128	131	125	137	137	129	116
171	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
152	172	165	97	113	125	122	129	165	133	176	98	128	111	173	201
117	155	161	167	149	143	159	116	93	137	116	198	122	152	127	203
116	114	118	116	154	201	134	133	99	123	123	131	171	121	119	127
110	108	134	162	129	119	137	141	109	125	149	103	111	238	172	154
119	127	116	114	118	116	123	121	154	201	134	133	99	123	123	131
159	116	93	137	116	198	122	152	127	203	117	155	161	167	149	143
123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
137	116	198	122	152	127	203	117	155	161	167	149	143	159	116	93
143	157	123	255	111	116	114	112	99	128	131	125	137	137	129	116
149	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
123	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
165	133	176	98	128	111	173	201	152	172	165	97	113	125	122	129
172	103	111	238	172	154	110	108	134	162	129	119	137	141	109	125
134	175	132	255	111	161	114	112	99	128	131	125	137	137	129	116
171	121	119	127	116	114	118	116	154	201	134	133	99	123	123	131
152	172	165	97	113	125	122	129	165	133	176	98	128	111	173	201
117	155	161	167	149	143	159	116	93	137	116	198	[illegible]	132	127	203
116	114	118	116	154	201	134	133	99	123	123	131	171	121	119	127
110	108	134	162	129	119	137	141	109	125	149	103	111	238	172	154
119	127	116	114	118	116	123	121	154	201	134	133	99	123	123	131
159	116	93	137	116	198	122	152	127	203	117	155	161	167	149	143

Topics in Computer Vision

1. Recognition
2. 3D Reconstruction
3. Motion Analysis
4. Low Level Vision

Logistics

Assessment

1. Assignments: 60%
2. Final Exam: 40%

Assignments

- Programming language: Python3
- Important libraries/frameworks: OpenCV
- Working environment: Jupyter notebook
- Submission is through Luminus
- Deadlines are strict:

  Late submission will be deducted 2 points (out of 10) for every 24 hours.

  Course website:
  https://tanrobby.github.io/teaching/ece_visual/index.html

Academic Honesty

Academic honesty is compulsory in finishing the assignments and the exam:

1. Exchanging codes is not allowed.
2. Using codes from the previous years or from the internet is prohibited, unless stated otherwise in the lectures.
3. Copying texts of the reports from other groups is strictly prohibited.
4. Generally, cheating, academic misconduct, plagiarism, and fabrication of any submitted material (including code and text) are not tolerated.

Any violation to the academic honesty will imply failure to pass the course.

References

- Textbooks (are used loosely):
  1. “Computer Vision: A Modern Approach”, D. Forsyth and J. Ponce
  2. "Computer Vision: Models, Learning and Inference”, S. Prince
  3. "Multiple View Geometry", R. Hartley and A. Zisserman

- Reading materials, slides, and lecture notes are available on the course website.

State-of-the-Art

- This module does NOT cover state-of-the-art. It teaches concepts needed to understand state-of-the-art
- For state-of-the-art:
  - Conference papers: CVPR, ICCV, ECCV, ICML, ICLR, NeurIPS, etc
  - Journal articles: IJCV, T-IP, JMLR, T-PAMI, etc
  - EE6733: Advanced Topics on Vision and Machine Learning
- EE5934/EE6934: Deep Learning

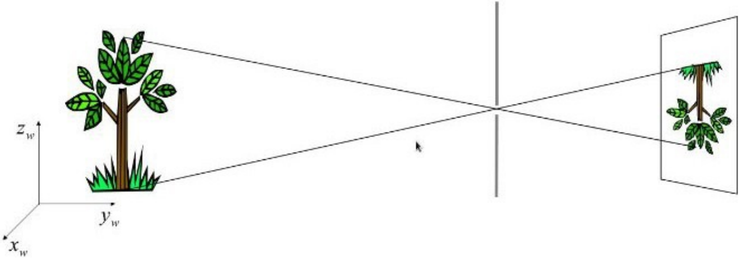
Image Formation

Computer Vision

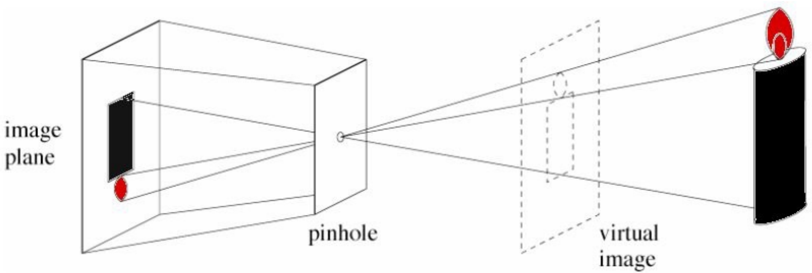
Robby T. Tan

Two fundamental aspects of image formation:

- Geometric Aspect
- Radiometric Aspect

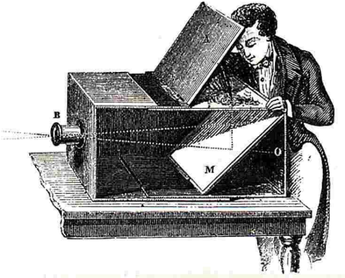


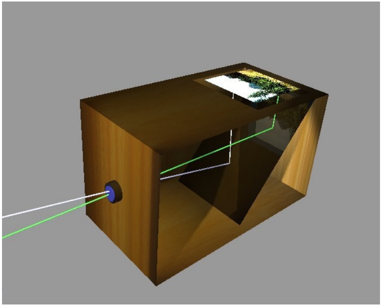
In the geometric perspective: an image is the projection of 3D structures onto a 2D image plane.



In a pinhole camera model, we can assume the presence of the virtual image.

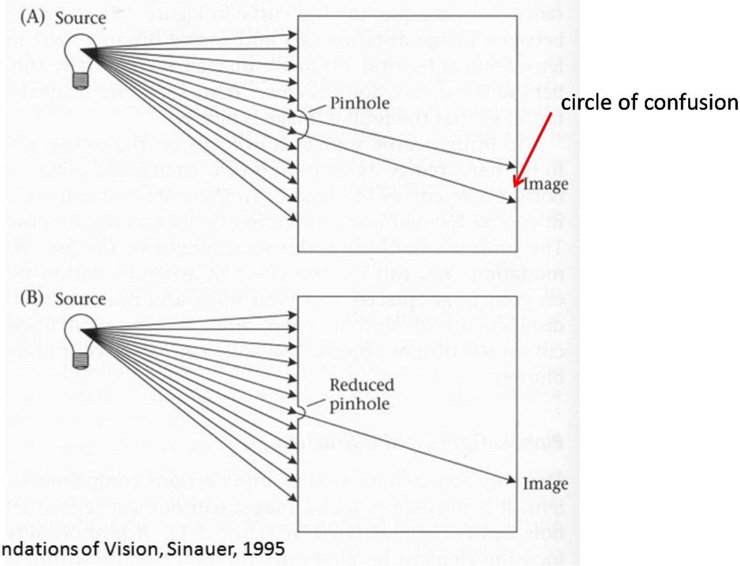
Pinhole Camera: Camera Obscura





Watch: https://youtu.be/gvzpu0Q9RTU

Aperture (Pinhole Size)



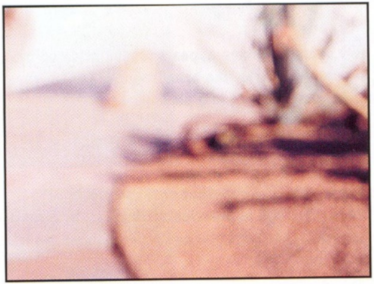
Wandell, Foundations of Vision, Sinauer, 1995

Aperture (Pinhole Size): Example

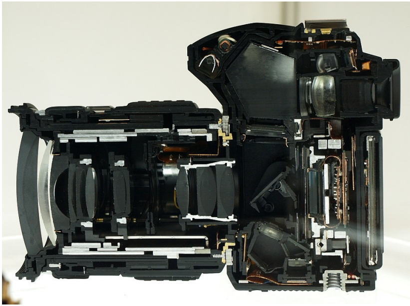
Photograph made with small pinhole

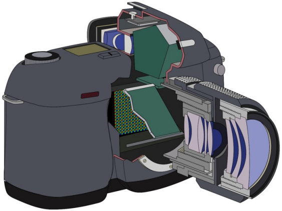
Sharp image

Photograph made with larger pinhole



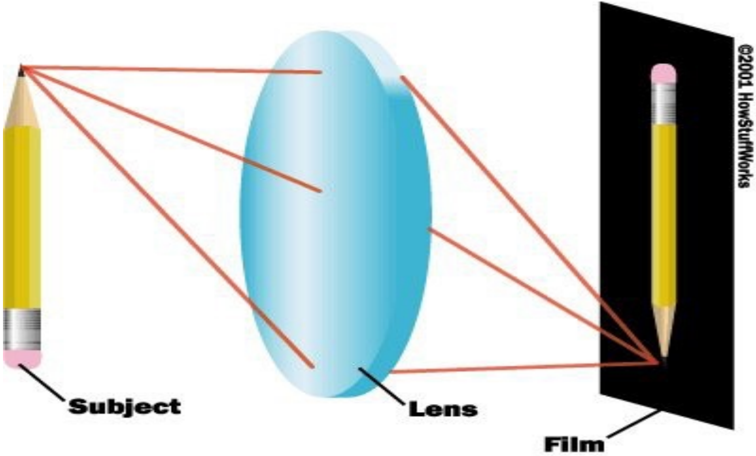
Blurry image





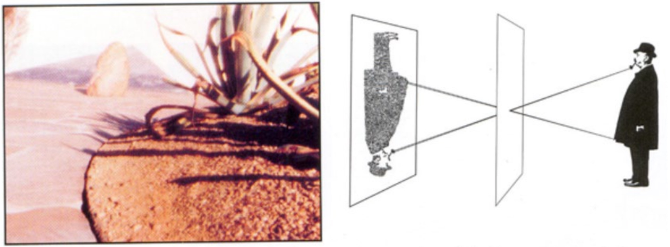
Modern cameras use lenses to have greater accumulated lights and different field of views

Digital Cameras: Lens

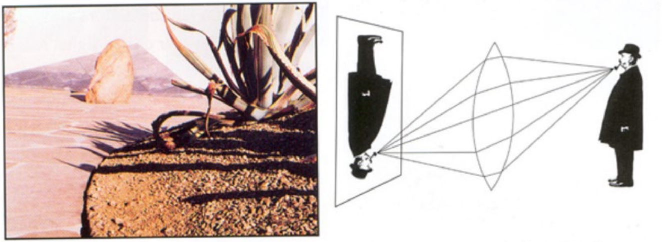


Lens: Benefit

Photograph made with small pinhole



Photograph made with lens



What are the geometric correlations between: a point in the 3D world, a pixel in the 2D image, and the camera?

The answer for this question will be discussed later.

Radiometric Aspect

- What is light?
- What is a digital image?
- How does a camera transform light into an image?

Electromagnetic Radiation

Light is electromagnetic radiation.

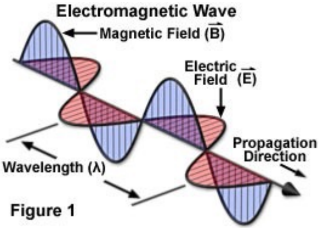
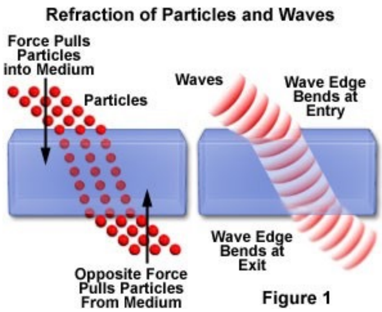


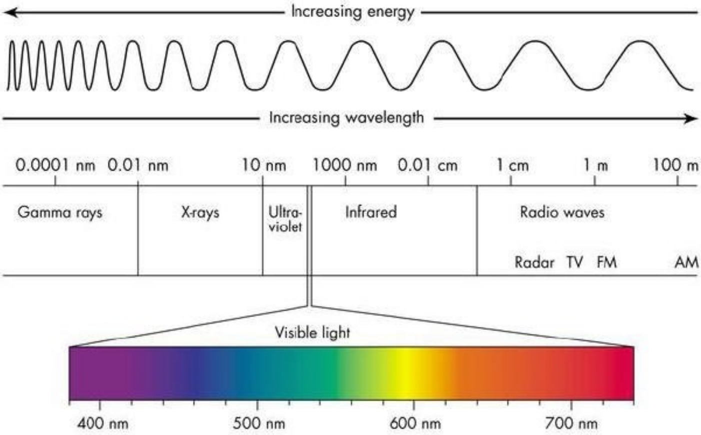
Figure 1

Wave-Particle Duality

Light is both wave and particle (wave-particle duality).



Lights as Waves: Wavelength



Lights as Waves: Light Sources

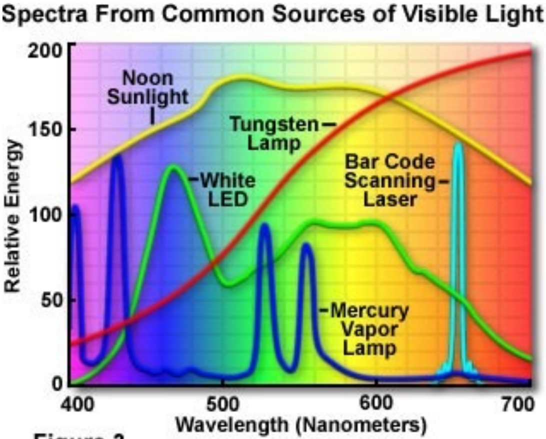
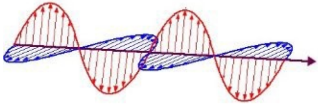


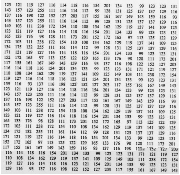
Figure 3



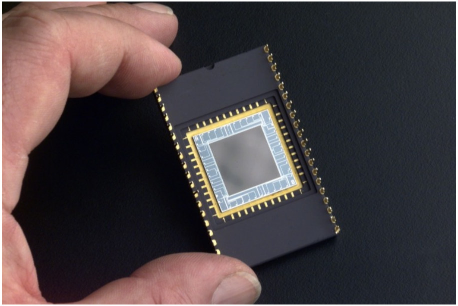
Digital camera: to convert light into electrical charges (then, to digital pixel intensities: $0 - 255$)

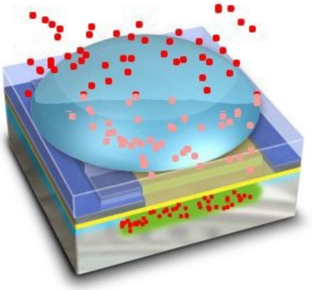
Basic components:

- Image sensors: CMOS or CCD
- Color Filters

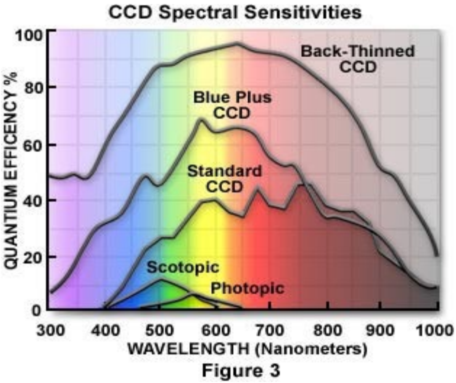


Charge Couple Devices (CCDs)



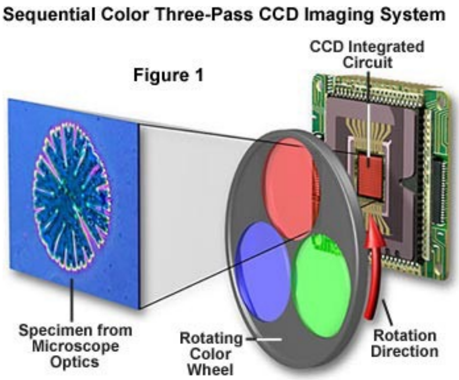


Spectral Sensitivity

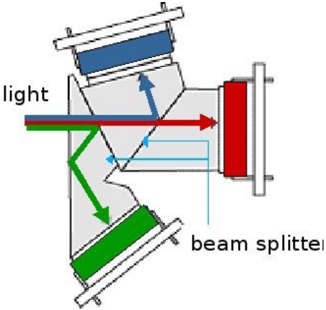


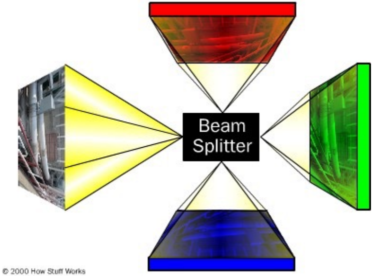
$$I(\mathbf{x}) = \int_{\Omega} E(\mathbf{x}, \lambda) q(\lambda) d\lambda \quad (1)$$

Color Filters: Color Wheel

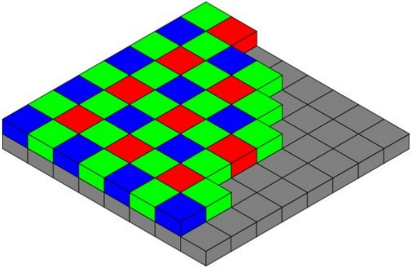


Color Filters: Beam Splitter





Color Filters: Bayer Filter



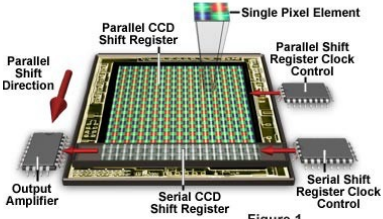
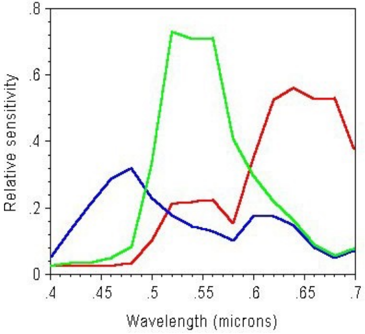
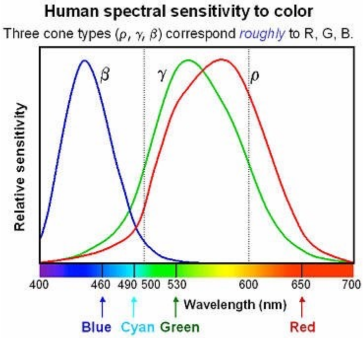


Figure 1

Different Cameras: Color Profiles

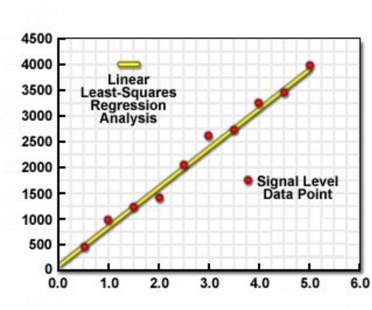
Spectral Sensitivity

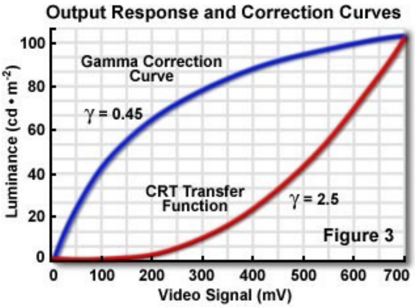




$$I_c(\mathbf{x}) = \int_{\Omega} E(\mathbf{x}, \lambda) q_c(\lambda) d\lambda \tag{2}$$

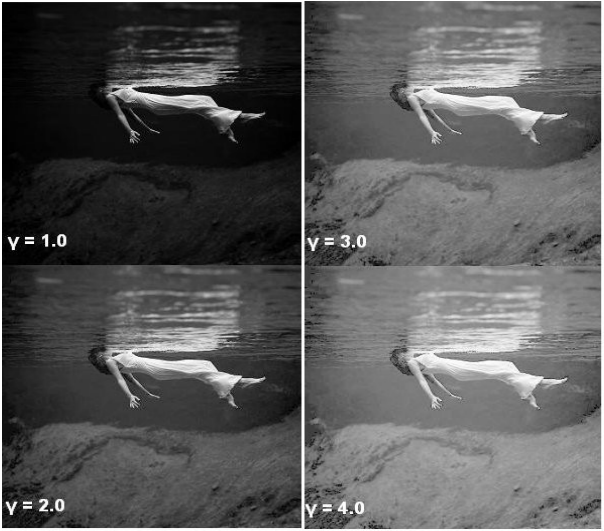
where index $c = \{R, G, B\}$





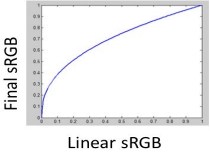
- The response function is determined by the camera manufacturers
- Radiometric calibration is to find the function and to linearize the response.

Why Non-linear?



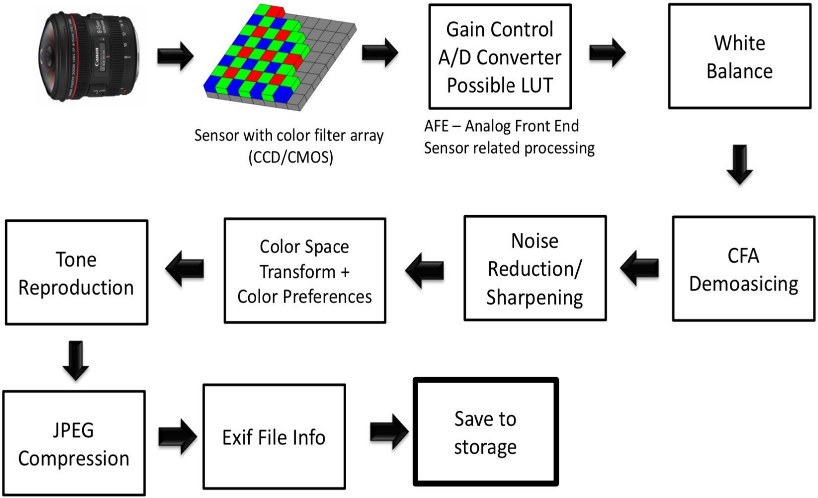
Gamma Function / Camera Response Function

Linear sRGB

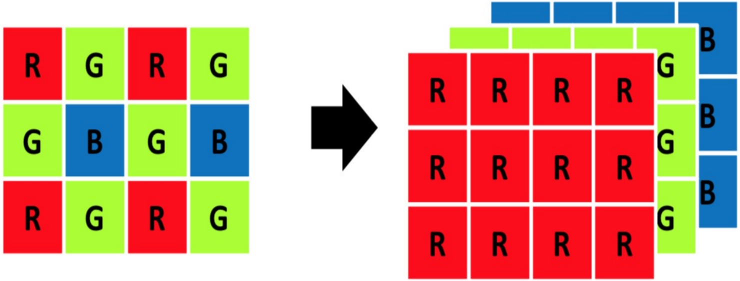


Final sRGB

Camera Pipeline



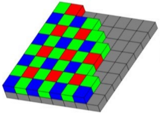
Camera Filter Array Demosaicing



Sensor RGB layout

Desired output with RGB per pixel.

RAW Image



Sensor with color filter array (CCD/CMOS)

Gain Control
A/D Converter
Possible LUT

AFE – Analog Front End
Sensor related processing

Lossless Compression

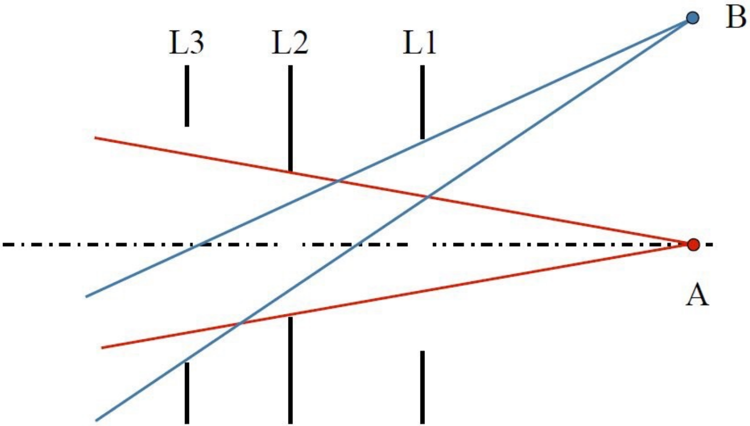
TIFF encoding (meta data)

Save to storage

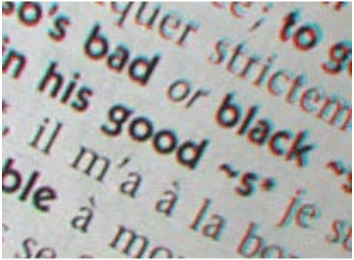
Camera Distortions

Vigneting

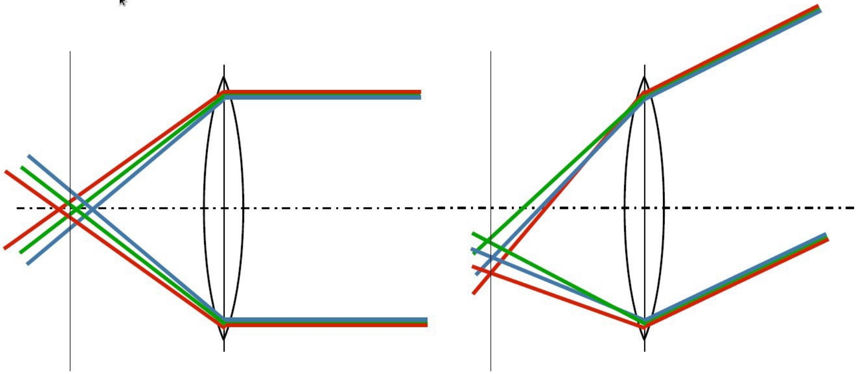
Cause of Vigneting



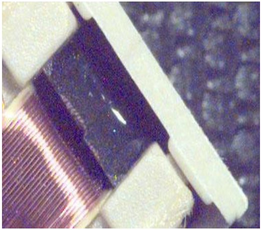
Chromatic Aberrations



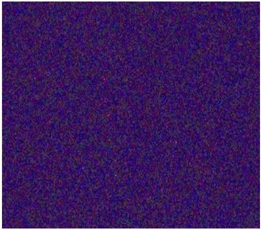
Cause of Aberrations



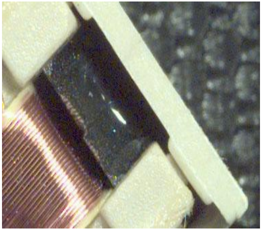
Dark Noise



image+dark noise

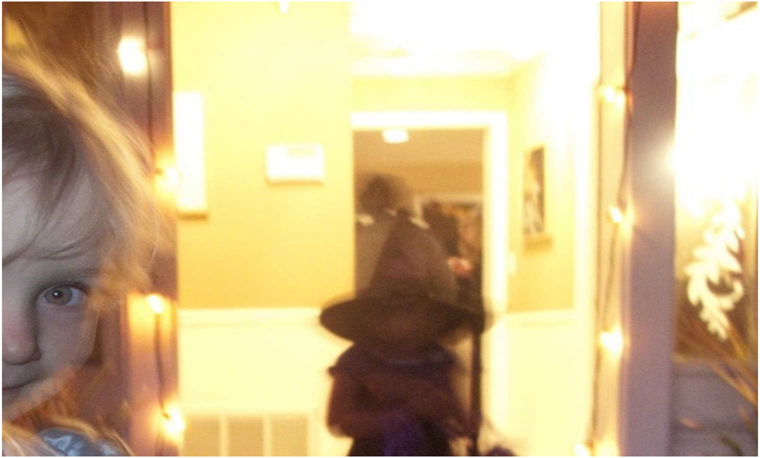


dark noise



subtraction result

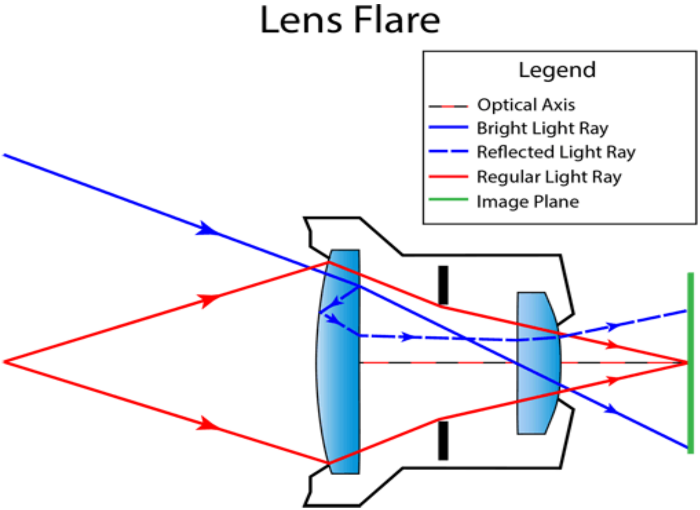
Pixel Saturation and Blooming



Lens Flare

Stray interreflection of light within the optical lens system.

Cause of Lens Flare



Stray interreflection of light within the optical lens system.

Geometric Radial Distortions

End of Slides